Maintaining representations of the environment of a mobile robot

Nicholas Ayache, Olivier Faugeras

INRIA

UNITÉ DE RECHERCHE
INRIA-ROCQUENCOURT

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P.105
78153 Le Chesnay Cedex
France
Tél.:(1) 39 63 55 11

Rapports de Recherche

N° 789

MAINTAINING REPRESENTATIONS OF THE ENVIRONMENT OF A MOBILE ROBOT

Nicholas AYACHE
Olivier D. FAUGERAS

FEVRIER 1988

International Symposium on Robotics Research, Santa-Cruz, California, August 1987

Maintaining Representations of the Environment of a Mobile Robot[1]

Nicholas Ayache and Olivier D. Faugeras

INRIA-Rocquencourt

BP 105 - 78153 Le Chesnay Cédex - France

Abstract: In this paper we describe our current ideas related to the problem of building and updating 3D representations of the environment of a mobile robot that uses passive Vision as its main sensory modality. Our basic tenet is that we want in these representations both geometry and uncertainty. We first motivate our approach by defining the problems we are trying to solve and giving some simple didactic examples. We then present the tool that we think is extremely well adapted to solving most of these problems: the Extended Kalman Filter (EKF). We discuss the notions of minimal geometric representations for 3D lines, planes, and rigid motions. We show how the EKF and the representations can be combined to provide solutions for some of the problems listed at the beginning of the article, and give a number of experimental results on real data.

Keywords: Fusion of Sensory Data, Mobile Robots, 3D Visual Maps, Geometric Representations, Uncertainty, Extended Kalman Filtering, Motion Estimation.

Mise à jour des représentations de l'environnement perçu par un robot mobile

Résumé: Nous décrivons l'état de nos recherches concernant la construction et la mise à jour de représentations 3D de l'environnement d'un robot mobile qui utilise la vision passive comme principal capteur. Les représentations utilisées incluent une description de la géométrie et de l'incertitude. Nous motivons d'abord notre approche par une liste des problèmes que nous désirons résoudre, illustrée par un ensemble d'exemples didactiques. Nous présentons ensuite un outil particulièrement bien adapté pour résoudre la plupart de ces problèmes: le Filtre de Kalman Etendu (FKE). Nous présentons les problèmes liés à la représentation minimale de primitives géométriques telles que les droites, les plans et les déplacements géométriques 3D. Nous montrons comment le FKE et ces représentations peuvent être combinées pour résoudre la plupart des poblèmes mentionnés au début de l'article, et nous présentons un ensemble de résultats expérimentaux obtenus sur des données réélles.

Mots-clés: *Intégration Multisensorielle, Robots Mobiles, Cartes visuelles 3D, Représentations Géométriques, Incertitude, Filtre de Kalman Etendu, Mouvement .*

---

[1]This work was partially supported by esprit project P940.

1 Introduction

In the last few years, Computer Vision has gone wildly into the area of 3D analysis from a variety of sensing modalities such as Stereo, Motion, Range Finders and sonars. A book that brings together some of this recent work is [Kan87].

Most of these sensing modalities start from pixels which are then converted into three-dimensional structures. A characteristic of this work as compared to previous work where images were the starting and the ending point (like in image restoration for example)is that noise in the measurements is of course still present but, contrary to what has happened in the past, it has to be taken into account all the way from pixels to 3D geometry.

Another aspect of the work on 3D follows from the observation that if noise is present, it has to evaluated, i.e. we need models of sensor noise (sensor being taken here in the broad sense of sensory modality), and reduced. This reduction can be obtained in many ways. The most important ones are:

- First, the case of one sensor in a fixed position: it can repeat its measurements and thus, maybe, obtain better estimations.
- Second, the case of a sensor that can be moved around: given its measurements in a given position, what is the best way to move in order to reduce the uncertainty and increase the knowledge of the environment in a way that is compatible with the task at hand.
- Third, is the case of several different sensors that have to combine their measurements in a meaningful fashion.

Interesting work related to these issues has already emerged which is not reported in [Kan87]. In the area of robust estimation procedures and models of sensors noise, Hager and Mintz [HM87] and Mc Kendall and Mintz [MM87] have started to pave the ground. Bolle and Cooper [BC86] have developed maximum likelyhood techniques to combine range data to estimate object positions. Durrant-Whyte [Dur86], in his Ph.D. Thesis has conducted a thorough investigation of the problems posed by Multi-Sensory systems. Applications to the navigation of a mobile robot have been discussed by Smith and Cheeseman [SC87] and by Matthies and Shafer [MS86]. The problem of combining stereo views has been attacked by Ayache and Faugeras [FAF86], [AF87b], Porril et al. [P*87], and Kriegman [KTT87].

Several problems related to these preliminary studies need more attention. Modelling sensor noise in general and more specifically visual sensor noise appears to us an area where considerable progress can be achieved ; relating sensor noise to geometric uncertainty and the corresponding problem of representing geometric information with an eye toward describing not only the geometry but also the uncertainty on this geometry are key problems to be investigated further as is the problem of combining uncertain geometric information produced by different sensors.

2 What are the problems that we are trying to solve

We have been focusing on a number of problems arising in connection with a robot moving in an indoor environment and using passive vision and proprioceptive sensory modalities such as odometry. Our mid-term goals are to incrementally build on the robot an increasing set of sensing and reasoning capabilities such as:

- build local 3D descriptions of the environment.
- use the descriptions to update or compute motion descriptions where the motion is either the robot's motion or others.
- fuse the local descriptions of neighboring places into more global, coherent, and accurate ones.
- "discover" interesting geometric relations in these descriptions.
- "discover" semantic entities and exhibit "intelligent" behavior.

We describe how we understand each of these capabilities and what are the underlying difficulties.

2.1 Build local 3D descriptions of the environment

Up until now, our main source of 3D information has been Stereo [AF87c,AL87] even though we have made considerable progress toward the use of structure from motion also [FLT87]. In any case the problems are very similar for both sensing modalities and we concentrate on Stereo. As announced in the introduction, our main concern is to track uncertainty all the way from pixel noise to geometric descriptions. Figure 1 shows for example that in a Stereo system, if pixels positions are imperfectly known, then the corresponding 3D point varies in an area with a quite anisotropic diamond shape. This is a clear example of a relation between pixel uncertainty and geometric (the position of point M) uncertainty. Another source of uncertainty in Stereo is the calibration uncertainty. In a stereo rig, intrinsic parameters of the cameras such as focal length, and extrinsic parameters such as relative position and orientation of the cameras have to be calculated. Figure 2 shows the effect on the reconstruction of a point M of an uncertainty on the focal lengths of the two cameras. Again, M varies in a diamond like shape. Of course this source of uncertainty adds itself to the previous pixel uncertainty.

Another example of the propagation of uncertainty is given in Figure 3 where pixels in left and right images are grouped into line segments: pixel uncertainty is converted into 2D line uncertainty. Line segments are then matched and used to reconstruct 3D line segments : 2D line uncertainty and calibration uncertainty are converted into 3D uncertainty. Yet another set of examples of this kind

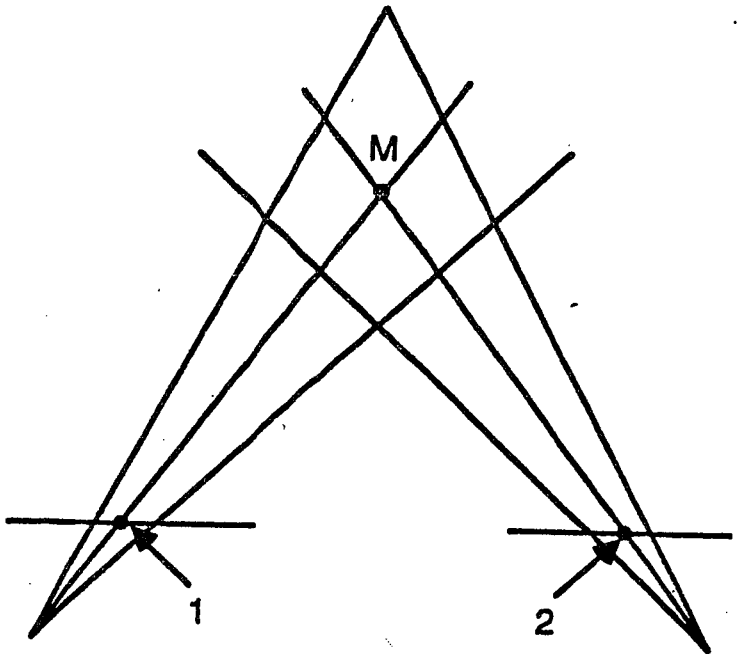

Figure 1: Effect of pixel noise on 3D reconstruction

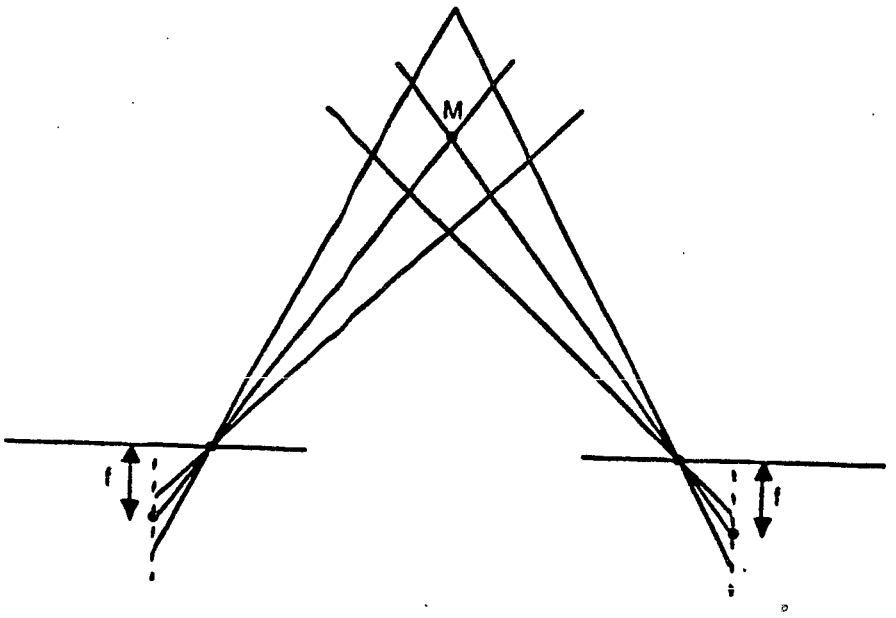

Figure 2: Effect of calibration errors on 3D reconstruction

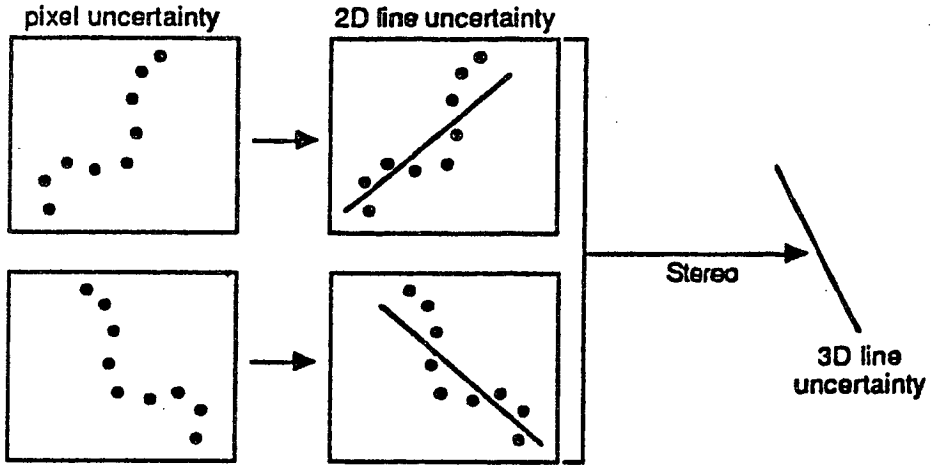

Figure 3: From pixel uncertainty to 3D line uncertainty

of propagation is shown in Figure 4 where coplanar and cocylindrical line segments are grouped together ; again, the question is, what is the uncertainty on the plane or on the cylinder ?

From these examples, we see that the main problem that needs to be solved in order to build local 3D descriptions of the environment is how geometric uncertainty propagates when we build up more complex primitives from simpler ones. This in turn generates two questions:

1. How do we represent geometric primitives.
2. How do we represent uncertainty on these primitives.

2.2 Update position and motion information

Figure 5 shows a measurement of a physical point made in two positions 1 and 2 of a mobile vehicle. In position 1, it "sees" M with some uncertainty represented by the ellipse around it. In position 2, it "sees" P with another uncertainty. Assuming that the displacement between 1 and 2 is exactly known, it is possible to express P and M in the same coordinate system. If the displacement estimate is wrong, as it is in Figure 5, the two zones of uncertainty do not intersect and it is very unlikely that the observer will realise that the points M and P are instances of the same physical point. If we now take into account the uncertainty on the displacement (assuming that we can estimate it) we have Figure 6 where the combination of displacement uncertainty and measurement uncertainty produces a larger ellipse around P which intersects the one around M : the observer can now infer that the probability of M and P being the same physical point is quite high and use the two measurements to obtain a better estimate of the displacement and reduce its uncertainty. The measurements can also be used to produce better estimates of the positions (Figure 7). This is related to what we call geometric fusion.

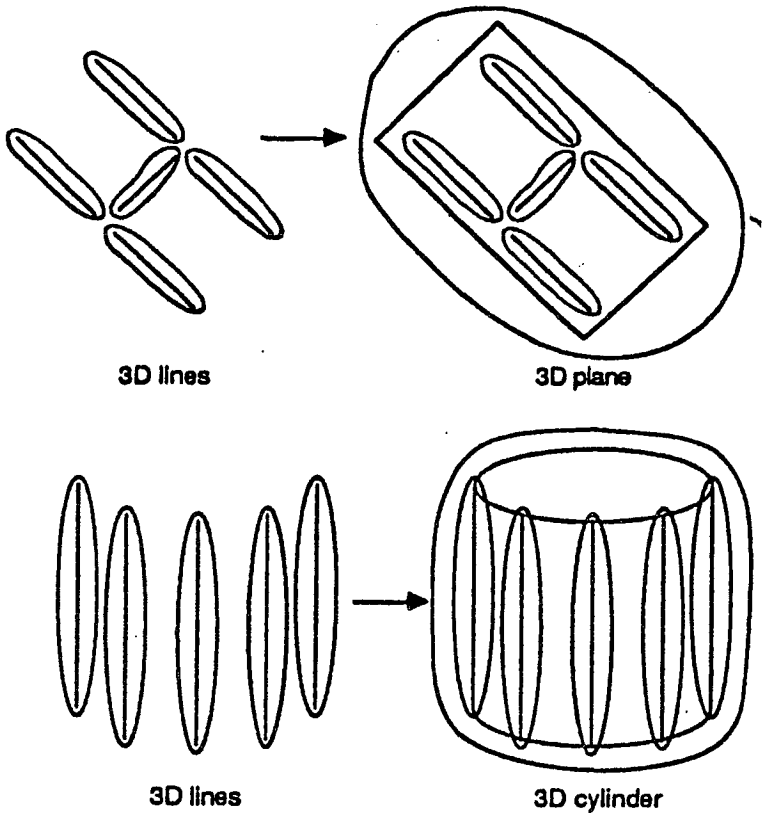

Figure 4: From 3D line uncertainty to 3D surface uncertainty

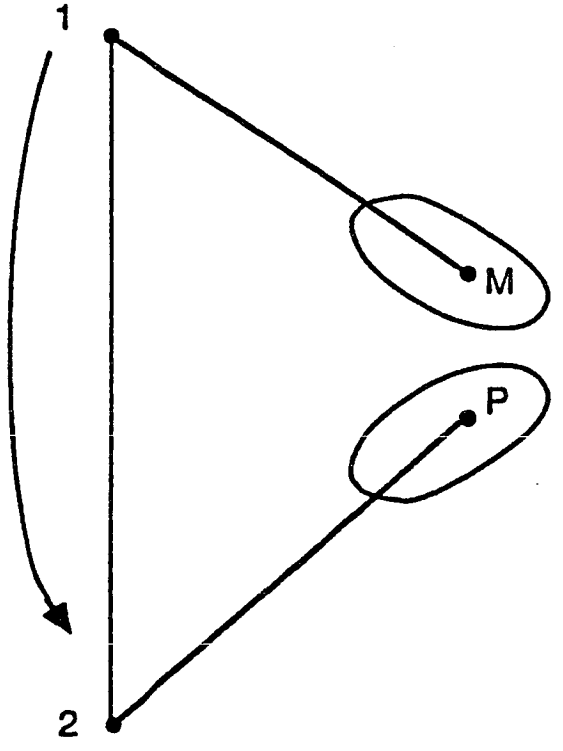

Figure 5: Measuring a point in two positions (wrong displacement estimation)

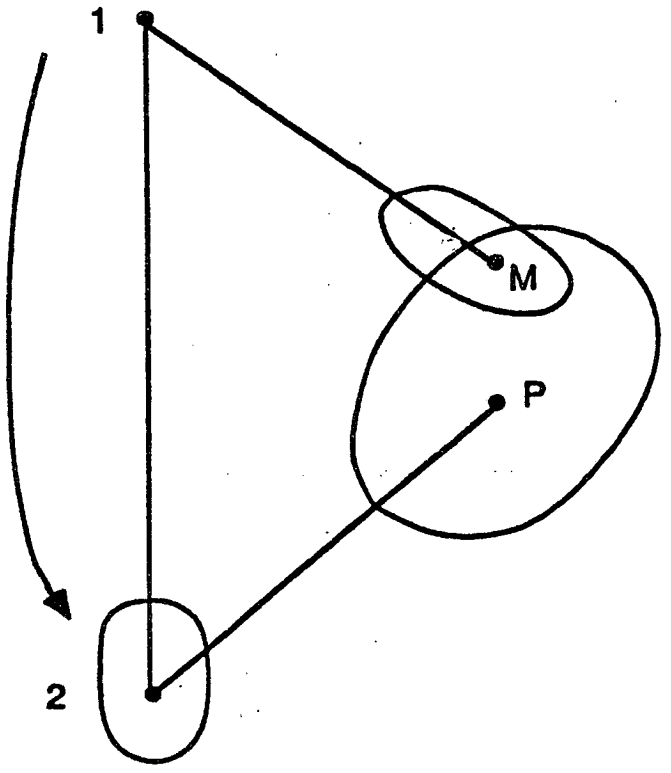

Figure 6: Measuring a point in two positions (displacement and uncertainty estimation)

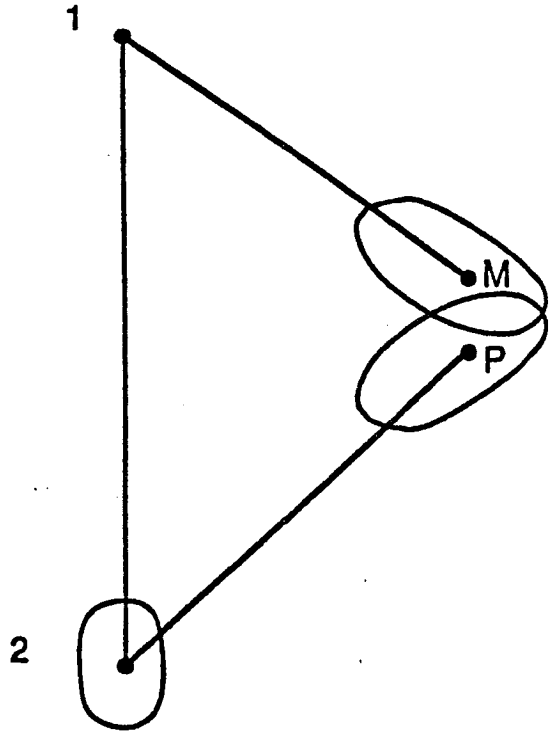

Figure 7: Improving the estimation of the points position

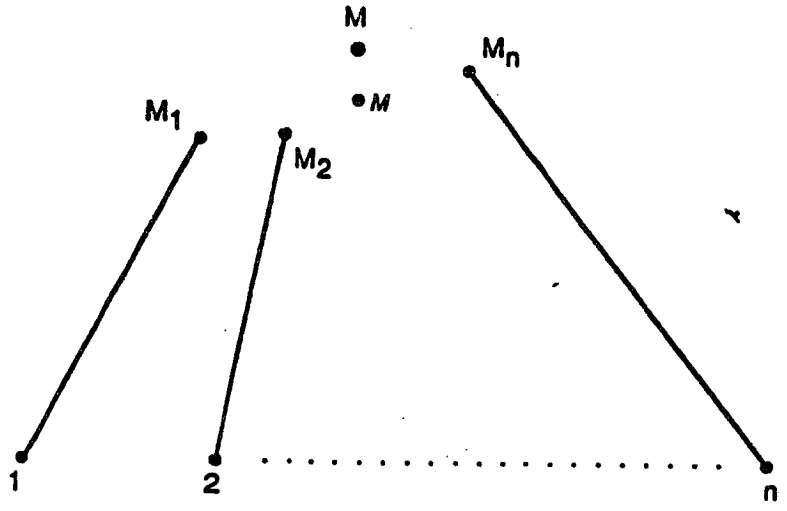

Figure 8: Fusing n points measured from different positions

2.3 Fusing geometric entities

Figure 8 shows a slightly more general case than what is depicted in Figure 7. The mobile vehicle has measured the physical point M in n positions numbered from 1 to n. Each measurement yields a point $M_i, i = 1, ..., n$ and some uncertainty in the coordinate system attached to the robot. Displacement uncertainty is also available. Using the ideas described in Section 5, we can improve the estimates of the displacements and reduce their uncertainty by discovering that points $M_1, ..., M_n$ are all instanciations of the same point. We can also use this observation to reduce the uncertainty on, let us say M_1, by combining the n measurements and produce a point $\mathcal{M}$, fusion of $M_1, ..., M_n$, as well as its related uncertainty. The points $M_1, ..., M_n$ can then be erased from the representation of the environment, they can be forgotten. What remains is the point $\mathcal{M}$ expressed in the coordinate system attached to position 1 for example, and the displacements from 1 to 2, 2 to 3, etc ..., which allow to express $\mathcal{M}$ in the other coordinate systems.

Fusing geometric entities is therefore the key to "intelligent" forgetting which in turn prevents the representation of the environment of growing too large.

2.4 Discovering "interesting" geometric relations

Using this approach also allows us to characterize the likelyhood that a given geometric relation exists between a number of geometric entities and to use this information to obtain better estimates of these entities and reduce their uncertainty. For example, as shown in Figure 9, segments AB and CD which have uncertainty attached to their endpoints have a high likelyhood to be parallel. Assuming that they are, we can update their position (they become more parallel) and reduce the uncertainty of their endpoints. The same reasoning can be used, for the relation "to be perpendicular".

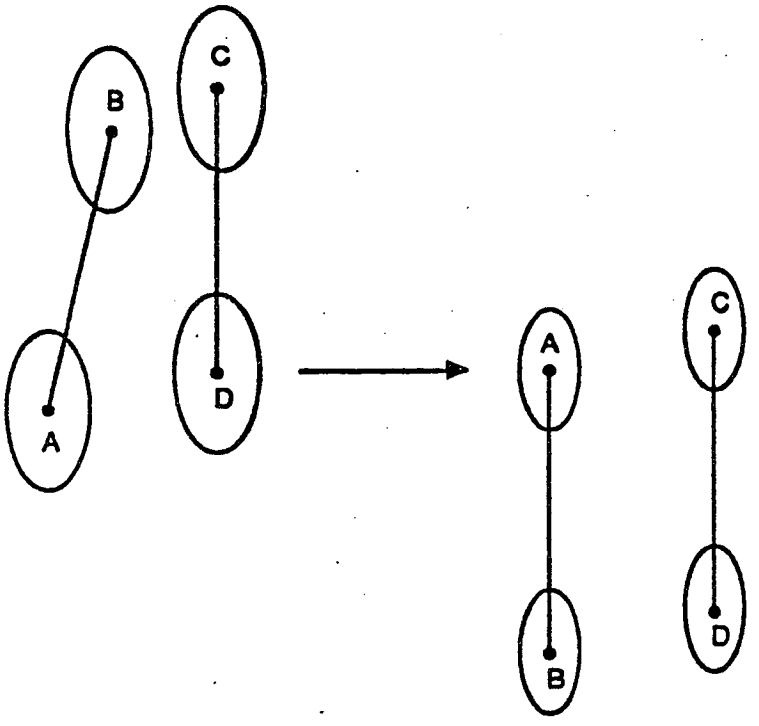

Figure 9: Discovering that AB and CD are parallel

2.5 Discovering semantic entities

Figure 10 shows the kind of "semantic" grouping that is of interest to us in the context of a mobile robot moving indoors, to combine geometry and some a priori description of the environment. The line segments numbered from 1 to 15 are found, using the ideas described in Section 2.4, to be coplanar with a high probability ; the corresponding plane is found to be vertical with a very high probability which can be deduced from the geometric uncertainty of the line segments. This observation can then be used to infer that the plane has a high probability to be a wall. If we also observe that segments 8 to 11 and 12 to 15 form approximately two rectangles this can be used to infer that they have a high probability to be parts of a window or a door.

3 What is the tool that we are using

In this third section, we introduce the Extended Kalman Filter (EKF) formalism which is applied in sections 4 and 5 to solve the problems we have just listed in section 2.

3.1 Unifying the problems

In all these previously listed problems, we are confronted with the estimation of an unknown parameter $a \in \mathbb{R}^n$ given a set of k non necessary linear equations of the form

$$f_i(x_i, a) = 0 \tag{1}$$

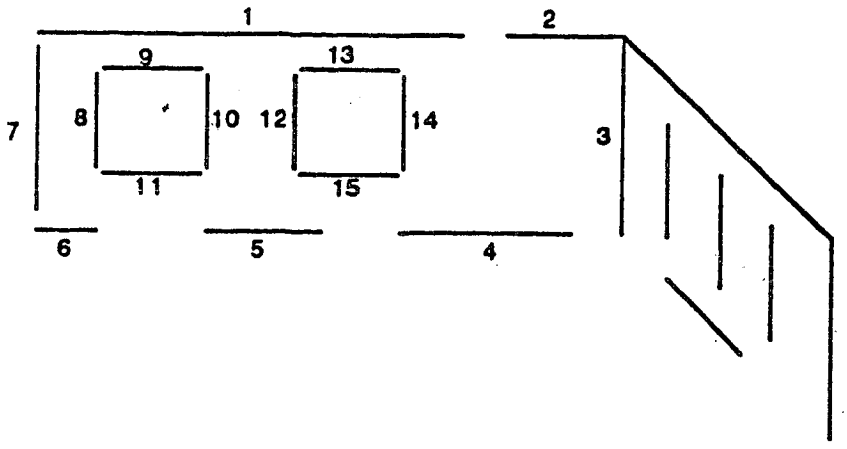

Figure 10: Hypothesizing walls, windows, and doors

where $x_i \in \mathbb{R}^m$ and f_i is a function from $\mathbb{R}^m \times \mathbb{R}^n$ into $\mathbb{R}^p$. The vector x_i represents some random parameters of the function f_i in the sense that we only measure an estimate $\hat{x}_i$ of them, such that

$$\hat{x}_i = x_i + v_i \tag{2}$$

where v_i is a random error. The only assumption we make on v_i is that its mean is zero, its covariance is known, and that it is a white noise :

$$E[v_i] = 0$$

$$E[v_i v_i^t] = \Lambda_i \geq 0$$

$$E[v_i v_j^t] = 0 \quad \forall\, i \neq j$$

These assumptions are reasonable. If the estimator is biased, it is possible to substract its mean to get an unbiased one. If we do not know the covariance of the error (or some other confidence measure on it), the estimator is meaningless. If two measurements $\hat{x}_i$ and $\hat{x}_j$ are correlated, we take the concatenation of them $\hat{x}_k = (\hat{x}_i, \hat{x}_j)$ and the concatenated vector function $f_k = [f_i^t, f_j^t]^t$. The problem is to find the optimal estimate $\hat{a}$ of a given the functions f_i and the measurements $\hat{x}_i$.

3.2 Linearizing the equations

The most powerful tools developed in parameter estimation are for linear systems. Before using complicated nonlinear optimization techniques, it is worthwile to try applying the linear tools to a linearized version of our equations. This is the EKF approach that we now develop.

For each nonlinear equation $f_i(x_i, a) = 0$ we need to know an estimate $\hat{a}_{i-1}$ of the solution a, and again a measure S_i of the confidence we have in this estimate[2]. We assume that $\hat{a}_{i-1}$ is given

[2]In practice, we shall see that only an initial estimate $(\hat{a}_0, S_0)$ of a is required prior to the first measurement $\hat{x}_1$, while the next ones $(\hat{a}_i, S_i)$ are provided automatically by the Kalman filter itself.

by

$$\hat{a}_{i-1} = a + w_i \tag{3}$$

where w_i is a random error. The only assumptions we make on w_i are the same as for v_i, i.e.

$$E[w_i] = 0$$

$$E[w_i\, w_i^t] = S_i \geq 0$$

where S_i is a given positive matrix. Here again, no assumption of gaussianness is required.

Having an estimate $\hat{a}_{i-1}$ of the solution, the equations are linearized by a first order Taylor expansion around $(\hat{x}_i, \hat{a}_{i-1})$:

$$f_i(x_i, a) = 0 \approx f_i(\hat{x}_i, \hat{a}_{i-1}) + \frac{\widehat{\partial f_i}}{\partial x}(x_i - \hat{x}_i) + \frac{\widehat{\partial f_i}}{\partial a}(a - \hat{a}_{i-1}) \tag{4}$$

where the derivatives $\widehat{\partial f_i}/\partial x$ and $\widehat{\partial f_i}/\partial a$ are estimated at $(\hat{x}_i, \hat{a}_{i-1})$:

Equation 4 can be rewritten as :

$$y_i = M_i a + u_i \tag{5}$$

where

$$y_i = -f_i(\hat{x}_i, \hat{a}_{i-1}) + \frac{\widehat{\partial f_i}}{\partial a}\hat{a}_{i-1}$$

$$M_i = \frac{\widehat{\partial f_i}}{\partial a}$$

$$u_i = \frac{\widehat{\partial f_i}}{\partial x}(x_i - \hat{x}_i)$$

Equation 5 is now a linear measurement equation, where y_i is the new measurement, M_i is the linear transformation, u_i is the random measurement error. Both y_i and M_i are readily computed from the actual measurement $\hat{x}_i$, the estimate $\hat{a}_{i-1}$ of a, the function f_i and its first derivative. The second-order statistics of u_i are derived easily from those of v_i :

$$E[u_i] = 0$$

$$W_i \triangleq E[u_i u_i^t] = \frac{\widehat{\partial f_i}}{\partial x}\Lambda_i \frac{\widehat{\partial f_i}^t}{\partial x}$$

3.3 Recursive Kalman Filter

When no gaussianness is assumed on the previous random errors u_i, v_i and w_i, the Kalman Filter equations provide the best (minimum variance) linear unbiased estimate of a. This means that among the estimators which seek $\hat{a}_k$ as a linear combination of the measurements $\{y_i\}$, it is the one which minimizes the expected error norm,

$$E[(\hat{a}_k - a)^t(\hat{a}_k - a)]$$

while verifying

$$E[\hat{a}_k] = a$$

The recursive equations of the Kalman Filter which provide a new estimate $(\hat{a}_i, S_i)$ of a from $(\hat{a}_{i-1}, S_{i-1})$ are the following ones [Jaz70]:

$$\hat{a}_i = \hat{a}_{i-1} + K_i(y_i - M_i\hat{a}_{i-1}) \tag{6}$$

$$K_i = S_{i-1}M_i^t(W_i + M_iS_{i-1}M_i^t)^{-1} \tag{7}$$

$$S_i = (I - K_iM_i)S_{i-1} \tag{8}$$

or equivalently

$$S_i^{-1} = S_{i-1}^{-1} + M_i^tW_i^{-1}M_i \tag{9}$$

One can see that the previously estimated parameter $\hat{a}_{i-1}$ is corrected by an amount proportional to the current error $y_i - M_i\hat{a}_{i-1}$ called the inovation. The proportionality factor, K_i, is called the Kalman gain. At the end of the process, $\hat{a}_k$ is the final estimate and S_k represents the covariance of the estimation error :

$$S_k = E[(\hat{a}_k - a)(\hat{a}_k - a)^t]$$

The recursive process is initialized by $\hat{a}_0$, an initial estimate of a, and S_0, its error covariance matrix. Actually, the criterion minimized by the final estimate $\hat{a}_k$ is :

$$C = (a - \hat{a}_0)^t\, S_0^{-1}(a - \hat{a}_0) + \sum_{i=1}^{k}(y_i - M_ia)^t\, W_i^{-1}(y_i - M_ia) \tag{10}$$

It is interesting to note that the first term of equation 10 measures the squared distance of a from an initial estimate, weighted by its covariance matrix, while the second term is nothing else than the classical least-square criterion, i.e. the sum of the squared measurement errors weighted by their covariance matrices. Indeed, intializing the process with an arbitrary $\hat{a}_0$ and $S_0^{-1} = 0$, criterion 10 provides the classical least-square estimate $\hat{a}_k$ obtained from the measurements only, while the initial estimate does not play any role.

The enormous advantage of such a recursive solution, is that if we decide, after a set of k measurements $\{x_i\}$, to stop the measures, we only have to keep $\hat{a}_k$ and S_k as the whole memory of the measurement process. If we decide later to take into account additional measurements, we simply have to initialize $\hat{a}_0 \equiv \hat{a}_k$ and $S_0 \equiv S_k$ and to process the new measurements to obtain exactly the same solution as if we had processed all the measurements together.

3.4 Gaussian assumption

Up to now, we did not introduce any gaussian assumption on the random measurement errors $v_i = x_i - \hat{x}_i$ of equation 2 and on the prior estimate error $w_0 = a - \hat{a}_0$ of equation 3. However, in practice, these errors usually come from a sum of independent random processes, which tend toward a gaussian process (Central Limit theorem). If we actually identify v_i and w_0 with gaussian processes, i.e.

$$v_i \equiv N(0, \Lambda_i)$$

$$w_0 \equiv N(0, S_0)$$

then, it follows that the noise u_i in equation 5 is also gaussian, i.e.

$$u_i \equiv N(0, W_i)$$

and that all the successive estimates provided by the recursive Kalman filter are also gaussian

$$\hat{a}_k = N(a, S_k)$$

Moreover, in this case, the Kalman filter provides the best (minimum variance) unbiased estimate $\hat{a}_k$ among all, even nonlinear, filters. This estimate $\hat{a}_k$ is also the maximum likelihood estimator of a. This comes from the fact that in the gaussian case, the solution is the conditional mean $\hat{a}_k = E[a/y_1, ..., y_k]$ which both minimizes the variance and maximizes the likelihood while being expressed as a linear combination of the measurements y_i. Therefore in this case, the minimum variance and minimum variance linear estimates are the same, namely the estimate $\hat{a}_k$ provided by the Kalman filter [Jaz70].

In conclusion, in the gaussian case, the Kalman filter provides the best estimate with the advantage of preserving gaussianness of all the implied random variables, which means that no information on the probability density functions of the parameters is lost while keeping only their mean and covariance matrix.

3.5 Rejecting Outlier Measurements

At iteration i, we have an estimate $\hat{a}_{i-1}$ and an attached covariance matrix S_{i-1} for parameter a. We also have a noisy measurement $(\hat{x}_i, \Lambda_i)$ of x_i and we want to test the plausibility of this mesurement with respect to the equation $f_i(x_i, a) = 0$.

If we consider again a first order expansion of $f_i(x_i, a)$ around $(\hat{x}_i, \hat{a}_{i-1})$ (equation 4), considering that $(\hat{x}_i - x_i)$ and $(\hat{a}_{i-1} - a)$ are independent centered gaussian processes, we see that $f_i(\hat{x}_i, \hat{a}_{i-1})$ is also (up to linear approximation) a centered gaussian process whose mean and covariance are given by:

$$E[f_i(\hat{x}_i, \hat{a}_{i-1})] = 0$$

$$Q_i = E[f_i(\hat{x}_i, \hat{a}_{i-1}) f_i(\hat{x}_i, \hat{a}_{i-1})^t] = \frac{\widehat{\partial f_i}}{\partial x} \Lambda_i \frac{\widehat{\partial f_i}^t}{\partial x} + \frac{\widehat{\partial f_i}}{\partial a} S_{i-1} \frac{\widehat{\partial f_i}^t}{\partial a}$$

Therefore, if the rank of Q_i is q, the generalized Mahalanobis distance :

$$d(\hat{x}_i, \hat{a}_{i-1}) = [f_i(\hat{x}_i, \hat{a}_{i-1})]^t Q_i^{-1} [f_i(\hat{x}_i, \hat{a}_{i-1})] \tag{11}$$

has a χ^2 distribution with q degrees of freedom [3].

Looking at a χ^2 distribution table, it is therefore possible to reject an outlier measurement $\hat{x}_i$ at a 95 % confidence rate by setting an appropriate threshold ϵ on the Mahalanobis distance, and by keeping only those measurements $\hat{x}_i$ which verify:

$$d(\hat{x}_i, \hat{a}_{i-1}) < \epsilon \tag{12}$$

3.6 Enforcing a perfect measurement

The recursive equations of the Kalman filter allow for the exact verification of a linear equation ([P*87]). This is done simply by setting the covariance of the linear measurement equation 5 to zero:

$$W_i = 0;$$

By doing this, it is easy to see that the Kalman gain becomes

$$K_i = S_{i-1} M_i^t (M_i S_{i-1} M_i^t)^{-1}$$

and that :

$$M_i K_i = I.$$

Therefore after processing the measurement, we have:

$$y_i - M_i \hat{a}_i = y_i - M_i \hat{a}_{i-1} - M_i K_i (y_i - M_i \hat{a}_{i-1}) = 0$$

which shows that the measurement equation 5 exactly holds. This fact yields a singularity of the covariance matrix, as can be seen by computing

$$S_i M_i^t = S_{i-1} M_i^t - S_{i-1} M_i^t (M_i S_{i-1} M_i^t)^{-1} M_i S_{i-1} M_i^t = 0$$

As was pointed out by Porrill [P*87], one can impose exact measurements only as long as they are independent. Before imposing a new exact measurement, one must verify that

$$S_{i-1} M_i^t \neq 0 \tag{13}$$

[3] If $q < p$ = the size of the measurement vector f_i, Q_i^{-1} is the pseudo-inverse of Q_i.

otherwise $M_i S_{i-1} M_i^t$ cannot be inverted. If $S_{i-1} M_i^t = 0$, it means that the new exact measurement is not independant of the previous ones. Either it verifies

$$y_i = M_i a_{i-1} \tag{14}$$

which means that it is already verified, or it is inconsistent with the previously enforced exact measurements and must be either rejected or considered as a noisy neasurement, i.e. with a covariance matrix $W_i > 0$. Of course in practice the tests of equations 13 and 14 are corrupted by numerical errors, and we did not test their sensitivity to such errors.

3.7 Iterated Extended Kalman Filter

Up to now we have found a powerful tool to deal with linear noisy systems of equations. But these linear equations come from the first order expansion of nonlinear equations. If the estimate $\hat{a}_i$ around which the expansion is performed is too far from the correct parameter a, the approximation of equation 4 is not very good, and the optimal solution of the linear system may differ significantly from the true one.

A method to reduce the effect of nonlinearities is to apply the iterated Kalman filter (cf. [Jaz70], pp. 279 and 353–355). This consists in applying equation 6 as long as $\hat{a}_i - \hat{a}_{i-1}$ is large enough, computing at each iteration a new value of K_i, y_i and M_i obtained from a re-linearization of f_i about the new estimate $\hat{a}_i$. In general, after a few iterations (especially at the begining of the process), $\hat{a}_i$ is so close to a that the linearization error is almost zero, yielding an almost optimal filter.

3.8 Extension to non-stationary process

We assumed in the previous sections that the parameter vector a of the measurement equations 1 was constant at each iteration i, i.e.

$$a_i = a_{i-1} \qquad \forall\, i$$

In fact, the Kalman filter naturally extends to non stationary processes for which the parameter vector a is the solution of a gauss-markov state equation at each iteration, i.e.

$$a_i = F_i a_{i-1} + r_i$$

where r_i is a gaussian centered white noise :

$$r_i \equiv N(0, R_i)$$

and

$$E[r_i r_j^t] = 0 \quad \forall\, i \neq j$$

In this case the Kalman equations are only slightly modified [Jaz70] but the general formalism remains the same.

This powerful extension of the Kalman filter has not been used up to now in our experiments. Nevertheless, it could provide elegant solutions in cases where the parameter a is non stationary, for instance to consider elastic deformations which are modeled by a slightly varying locally rigid transformation.

4 Geometric Representations

In this Section, we give the details of the geometric representations that we have found useful at various stages of our work. It is first important to note that we have been dealing so far only with points, lines, and planes, i.e. with affine geometric entities. This may appear to be quite a restriction on the type of environments that we can cope with. This is indeed the case but there are a number of reasons why we think that our approach is quite reasonable:

1. The obvious one is that for the kind of environment that our mobile robot moves into, these primitives are very likely to cover most of the geometric features of importance.
2. A second reason is that more complicated curved features can be first approximated with affine primitives which are then grouped into more complicated non affine primitives.
3. A third reason is that we believe that the techniques we have developed for representing and combining uncertainty of affine primitives are generic and directly applicable to non affine primitives.

Let us now discuss specifically lines, planes, and rigid displacements.

4.1 Line segments

The 3D segments that we deal with are usually constructed from stereo [AF87c,AL87]. Their endpoints may be quite unreliable, even though they can be of some usefulness from time to time, and we largely depend on the infinite lines supporting those line segments.

We concentrate here on how to represent 3D lines. The obvious representation we mention here only for pedagogical reasons, is by two points; this representation is 6-dimensional and, as we will see next, not minimal. Another way to represent a line is to choose a point on it (3 parameters), and a unit vector defining its direction (2 parameters). The corresponding representation is 5-dimensional and, again, not minimal. In fact, the set of affine 3D lines is a manifold of dimension 4 for which we will exhibit later an atlas of class C^∞. This implies that a minimal representation of a straight line has four parameters. One such representation can be obtained by considering the normal to the line from the origin (if the line goes through the origin it is the same as a vector line

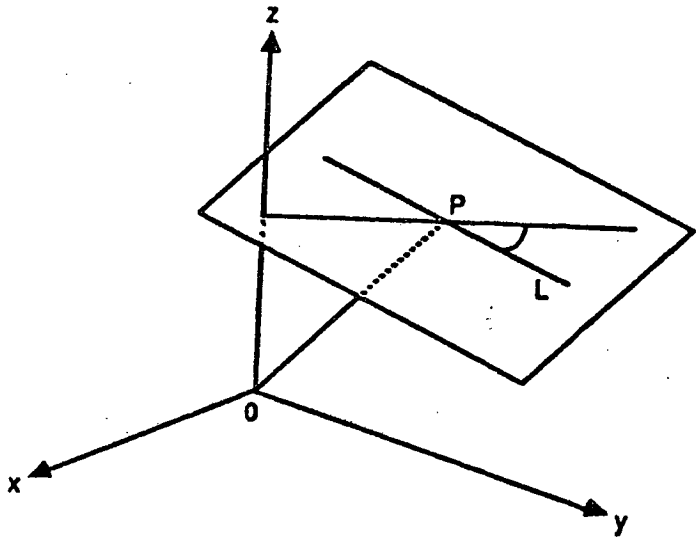

Figure 11: 3D line representation

and can be defined by two parameters only). The point of intersection between the normal and the line is represented by three parameters. If we now consider (see Figure 11) the plane normal at P to OP, the line is in that plane and can be defined by one more parameter, its angle with an arbitrary direction, for example the line defined by P and one of the axis of coordinates (in Figure 11, the z-axis). Of course, when the line is parallel to the xy-plane this direction is not defined and we must use either the x- or y-axis. This brings up an interesting point, namely that there does not exist a global minimal representation for affine lines, i.e. one which can be used for all such lines. We must choose the representation as a function of the line orientation. Mathematically, this means that the manifold of the affine straight lines cannot be defined with only one map. This is quite common and is also true for affine planes and rotations of $\mathbb{R}^3$, as will be shown next.

The previous representation for a line is not in fact the one we have been using. The one we use considers a line (not perpendicular to the z-axis) as the intersection of a plane parallel to the y-axis, and a plane parallel to the x-axis:

$$\begin{cases} x = az + p \\ y = bz + q \end{cases} \tag{15}$$

The intersection is represented by the four-dimensional vector $L = [a, b, p, q]^T$ which has the following geometric interpretation. The direction of the line is that of the vector $[a, b, 1]^T$, and the point of intersection of the line with the xy-plane has coordinates p and q. Since the last coordinate of the direction vector is equal to 1, the line cannot be perpendicular to the z-axis or parallel to the xy-plane. If we want, and we do in pratice, represent such lines, we must choose another

representation, for example:

$$\begin{cases} y = ax + p \\ z = bx + q \end{cases} \tag{16}$$

which cannot represent lines parallel to the yz-plane, or perpendicular to the x-axis, or:

$$\begin{cases} z = ay + p \\ x = by + q \end{cases} \tag{17}$$

which excludes lines parallel to the zx-plane.

Each representation defines a one to one mapping between $\mathbb{R}^4$ and a subset (in fact an open subset) of the set of affine 3D lines and it can be shown that these three mapping define on this set a structure of C^∞-manifold for which they form an atlas.

In practice, this means the representation is not exactly four-dimensional, but is made of the four numbers a, b, p, and q and an integer i taking the values 1, 2, and 3 to indicate which map 15, 16, or 17 we are currently using.

The fact that the set of affine 3D lines has been given a structure of C^∞-manifold implies that the a', b', p', q' of a given representation are C^∞-functions of the a, b, p, q of another representation for all lines for which the two representations are well defined (for example all lines not parallel to the xy- and yz-planes).

The representation of a line also includes a 4×4 covariance matrix Λ_L on the vector L. It is interesting at this stage to trace the computation of this covariance matrix all the way from pixel to 3D. In order to do this, we must briefly explain how 3D lines are computed in our current Stereo system [AL87]. We use three cameras as indicated in Figure 12. In theory, the three planes defined by the 2D lines l_1, l_2, and l_3 and the optical centers C_1, C_2, and C_3 belong to the same pencil and intersect along the 3D line L. In practice they do not because of noise and we have to find the "best" line satisfying the measurements, i.e. l_1, l_2 and l_3. This can be done by using the idea of pencil of planes, described more fully in [FLT87]. We assume that in the coordinate system attached to camera 1, for example, the equation of the ith plane P_i, $i = 1, 2, 3$, is given by

$$u_i x + v_i y + w_i z + r_i = 0$$

where the 4 vectors $P_i = [u_i, v_i, w_i, r_i]^T$ are known, as well as their covariance matrix Λ_{P_i} (we show later how to compute them). If we use representation 15 for the 3D line, it is represented as the intersection of the two planes P of equation $x = az + p$ and Q of equation $y = bz + q$. Writing that the five planes P, Q, and P_i, $i = 1, 2, 3$, from a pencil allows us to write 6 equations :

$$\begin{cases} w_i + au_i + bv_i = 0 \\ r_i + pu_i + qu_i = 0 \end{cases} \quad i = 1, 2, 3.$$

in the four unknowns a, b, p, and q.

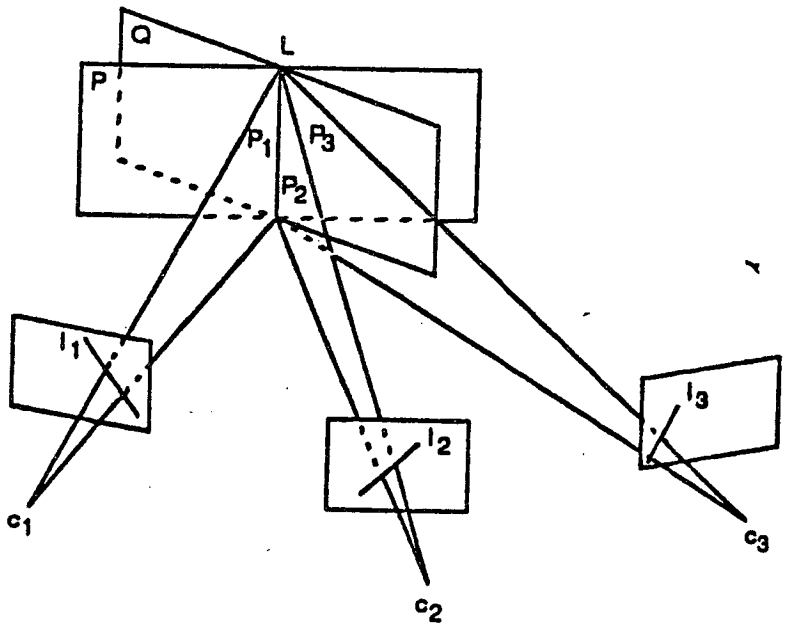

Figure 12: Reconstruction of 3D lines

We can apply directly the Kalman formalism to these measurement equations and choose $a = [a, b, p, q]^T$, and x_i as the 4-vector P_i. We can therefore simply compute an estimate $\hat{a}$ of a and its covariance matrix $\Lambda_{\hat{a}}$ from the P_i's and Λ_{P_i}'s.

Let us now show how we can compute the P_i's and Λ_{P_i}'s. Each line $l_i, i = 1, 2, 3$, is obtained by fitting a straight line to a set of edge pixels which have been detected using a modified version of the Canny edge detector [Can86,Der87]. Looking at Figure 13, let $xcos\theta + ysin\theta - p = 0$ be the equation of the line l which is fit to the edge pixels m_i of coordinates x_i, y_i ($0 \leq \theta < 2\pi, \rho \geq 0$). We assume that the measured edge pixels are independent and corrupted by a gaussian isotropic noise and take the parameter a equal to $[\theta, \rho]^T$ and the measurement x as the vector $[x, y]^T$. The measurement equation is therefore :

$$f(x, a) = xcos\theta + ysin\theta - \rho$$

Applying the EKF formalism to the n edge pixels forming the line provides the best estimate $\hat{a}$ of the line parameters and its covariance matrix. Having done this for all three cameras, it is easy to deduce the equations of the three planes P_i and the covariance matrixes on their coefficients.

4.2 Planes

Planes can receive pretty much the same treatment as lines. A plane is defined by three parameters, and this is minimal. A possible representation is the representation by the normal n (a unit norm vector), and the distance d to the origin. The problem with this representation is that it is not unique since $(-n, -d)$ represents the same plane. It is possible to fix that problem by assuming

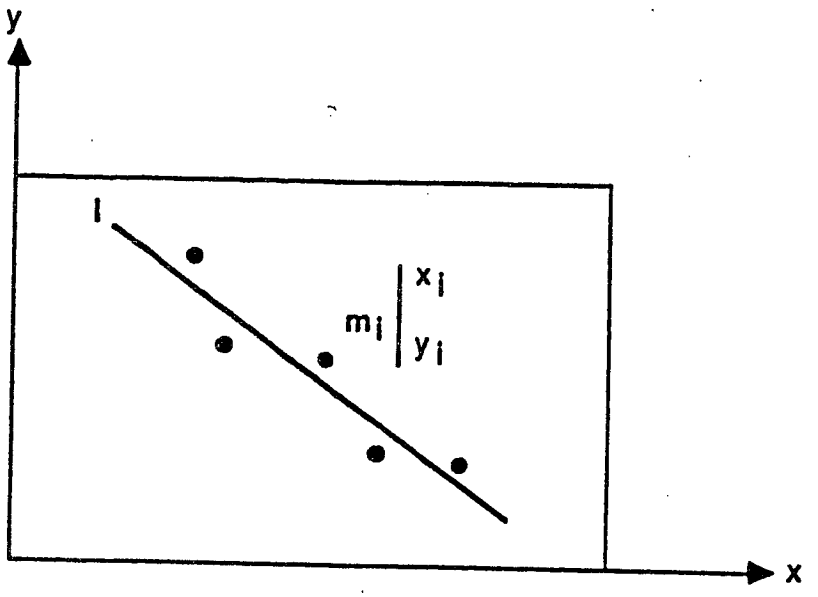

Figure 13: 2D line approximation

that one component of n, let us say n_z, is positive, i.e. we consider planes not parallel to the z-axis. For these planes we must choose another convention, for example that n_x is positive. Again, this works well for planes not parallel to the x-axis. The third possible representation is to assume n_y positive which excludes planes parallel to the y-axis.

So, we have three one to one mappings of open subsets of the product $S_2 \times R$, where S_2 is the usual gaussian sphere, into open subsets of the set of planes :

$$(n,d),\ n_z > 0 \longrightarrow \text{planes non parallel to } Oz$$
$$(n,d),\ n_x > 0 \longrightarrow \text{planes non parallel to } Ox$$
$$(n,d),\ n_y > 0 \longrightarrow \text{planes non parallel to } Oy$$

It is easy to show that these three mappings define on the set of 3D planes a structure of C^∞-manifold of dimension 3.

One practical disadvantage of the previous representations is that the normal n is constrained to lie on the unit sphere S_2 i.e. must satisfy the constraint $\|n\| = 1$. A possibly simpler representation is obtained by considering the mapping from R^3 to the set of 3D planes defined by :

$$p_1 : (a,b,c) \longrightarrow ax + by + z + c = 0 \tag{18}$$

this can represent all planes except those parallel to Oz and it is a one to one continuous mapping from $\mathbb{R}^3$ to the open subset of the set of 3D planes constituted of the planes not parallel to the z-axis. In order to obtain all possible plane, we must also consider the mappings :

$$p_2 : (a,b,c) \longrightarrow x + ay + bz + c = 0 \tag{19}$$
$$p_3 : (a,b,c) \longrightarrow bx + y + az + c = 0 \tag{20}$$

p_2 (resp. p_3) excludes planes parallel to the x-axis (resp. the y-axis). It is easy to show that p_1, p_2, p_3 also define on the set of 3D planes a structure of C^∞-manifold of dimension 3.

4.3 Rigid displacements

In a previous papers [AF87b] we have proposed to use the exponential representation of rotation matrixes. This is the same as saying that a rotation is defined by its axis u (a unit vector) and its angle θ. The vector $r = \theta u$ can be used to represent the rotation and we have :

$$R = e^H$$

where H is an antisymmetric matrix representing the cross-product with the vector r (i.e. $Hx = r \wedge x$, for all x).

Let us see how we can define a structure of manifold on the set of rotations using this representation. If we allow θ to vary over the semi-open interval $[0, 2\pi[$, the vector r can vary in the open ball $B(0, 2\pi)$ of $\mathbb{R}^3$ of radius 2π. But the mapping $f : B(0, 2\pi)$ into the set of rotations is not one to one because (u, π) and $(-u, \pi)$ represent the same rotation. To enforce uniqueness we can assume that one of the coordinates, for example u_z, of the rotation axis u is positive. We can then represent uniquely the open subset of the set of rotations for which the axis is not perpendicular to the z-axis, and has a positive component along that axis, and the mapping is continuous. If we consider the open set of rotations defined by $(u, \theta), u_z < 0$, we have another one to one continuous mapping. With these two mappings, we cannot represent rotations with an axis perpendicular to the z-axis. In order to obtain all possible rotations, we have to introduce the other four mappings defined by (u, θ) and $u_x > 0$ (resp. $u_x < 0$, $u_y > 0$, $u_y < 0$) which represent rotations with an axis not perpendicular to the x-axis (resp. the y-axis). We are still missing the null vector, i.e we have no representation for the null rotation, the identity matrix. In order to include it, we have to add a seventh map by considering for example the rotations defined by the "small" open ball $B(0, \epsilon)$ where ϵ must be smaller than π. These seven mappings define on the set of rotations a structure of C^∞-manifold of dimension 3.

It is interesting that in all three cases (3D lines, planes, and rotations), there does not exist a unique global representation and that we must live with at least three local mappings.

It is now instructive to study how the group of rigid displacements operates on the representations for lines and planes.

4.3.1 Lines

The easiest way to derive how representation 15 changes under rotation and translation is by considering that the line is defined by two points M_1 and M_2 of coordinates (x_1, y_1, z_1) and (x_2, y_2, z_2).

It is then easy to verify that:

$$a = \frac{x_2 - x_1}{z_2 - z_1} \qquad b = \frac{y_2 - y_1}{z_2 - z_1}$$
$$p = \frac{z_2\, x_1 - z_1\, x_2}{z_2 - z_1} \qquad q = \frac{z_2\, y_1 - z_1\, y_2}{z_2 - z_1}$$

Introducing the vector $M_1M_2 = [A, B, C]^T$, we have $a = A/C$, and $b = B/C$. a, and b are therefore only sensitive to rotation:

$$M_1M_2 \longrightarrow R\, M_1M_2$$
$$\begin{bmatrix} A \\ B \\ C \end{bmatrix} \rightarrow \begin{bmatrix} A' \\ B' \\ C' \end{bmatrix} = R \begin{bmatrix} A \\ B \\ C \end{bmatrix}$$
$$\begin{bmatrix} a \\ b \end{bmatrix} \rightarrow \begin{bmatrix} a' \\ b' \end{bmatrix} = \begin{bmatrix} A'/C' \\ B'/C' \end{bmatrix}$$

This yields :

$$a' = \frac{r_1.\alpha}{r_3.\alpha} \qquad b' = \frac{r_2.\alpha}{r_3.\alpha}$$

where $\alpha = [a, b, 1]^T$, and the r_i's are the row vectors of matrix R. This is true only if $r_3.\alpha \neq 0$; if $r_3.\alpha = 0$, the transformed line is perpendicular to the z-axis and representation 16 or 17 must be used.

To treat the case of p and q, let us introduce $P = p(z_2 - z_1) = pC$ and $Q = q(z_2 - z_1) = qC$. It is easy to show that :

$$\begin{bmatrix} P \\ Q \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix} OM_1 \wedge OM_2 = H(OM_1 \wedge OM_2)$$

This allows us to study how P and Q change under rotation and translation :

$$OM_1 \rightarrow R\,OM_1 + t \quad OM_2 \rightarrow R\,OM_2 + t$$

therefore :

$$\begin{bmatrix} P \\ Q \end{bmatrix} \rightarrow \begin{bmatrix} P' \\ Q' \end{bmatrix} = H(R(OM_1 \wedge OM_2) + t \wedge R\, M_1M_2)$$

Using the previous notations, $M_1M_2 = [A, B, C]^T$, and $OM_1 \wedge OM_2 = [Q, -P, X]^T$ where X is unknown. But noticing that $M_1M_2.(OM_1 \wedge OM_2) = 0$ we have :

$$AQ - BP + CX = 0$$

and therefore :

$$X = \frac{BP - AQ}{C} = bP - aQ$$

C is not equal to 0 since by definition, the line is not perpendicular to the z-axis.Putting everything together:

$$\begin{bmatrix} P' \\ Q' \end{bmatrix} = CH \left(R \begin{pmatrix} q \\ -p \\ bp - aq \end{pmatrix} + t \wedge R\alpha \right)$$

Finally :

$$\begin{bmatrix} p' \\ q' \end{bmatrix} = \begin{bmatrix} P'/C' \\ Q'/C' \end{bmatrix} = \frac{C}{C'} H \left(R \begin{pmatrix} q \\ -p \\ bp - aq \end{pmatrix} + t \wedge R\alpha \right)$$

and we know from the previous derivation that $C/C' = 1/r_3.\alpha$, therefore :

$$\begin{bmatrix} p' \\ q' \end{bmatrix} = \frac{1}{r_3.\alpha} H(R\beta + t \wedge R\alpha)$$

where we have taken $\beta = [q, -p, bp - aq]^T$.

4.3.2 Planes

Given a plane represented by its normal n and its distance to the origin d, if we apply to it a rotation along an axis going through the origin represented by a matrix R followed by a translation represented by a vector t, the new plane is represented by Rn and $d - t.Rn$ [FH86].

This allows us to compute how the representation 18, for example, is transformed by the rigid displacement. From the previous observation :

$$\begin{pmatrix} a \\ b \\ 1 \end{pmatrix} \rightarrow R \begin{pmatrix} a \\ b \\ 1 \end{pmatrix} \text{ and } c \rightarrow c - t.R \begin{pmatrix} a \\ b \\ 1 \end{pmatrix}$$

Introducing the three row vectors r_1, r_2, r_3 of matrix R, we have, assuming that $r_3.\alpha \neq 0$:

$$a' = \frac{r_1.\alpha}{r_3.\alpha} \quad b' = \frac{r_2.\alpha}{r_3.\alpha} \quad c' = \frac{c - t.R\alpha}{r_3.\alpha}$$

if $r_3.\alpha = 0$, this means that we cannot use the same representation for the transformed plane since it is parallel to the z-axis, therefore we must choose the representation 19 or 20.

5 Registration, Motion and Fusion of Visual Maps

In this section we show how to solve the problems listed in section 2 within the formalism and the representations explicited in sections 3 and 4.

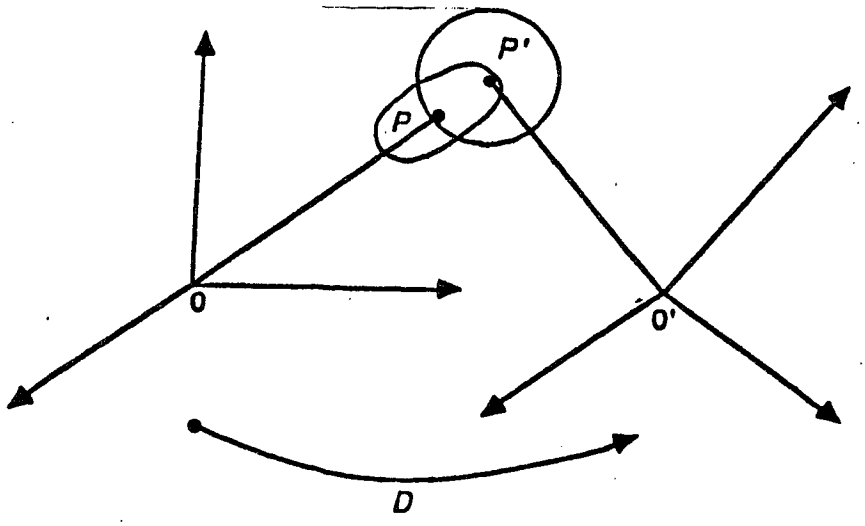

Figure 14: The general registration motion fusion problem

5.1 Initial assumptions

We are given two visual maps $\mathcal{V}$ and $\mathcal{V}'$, each of them attached to a coordinate reference frame $\mathcal{F}$ and $\mathcal{F}'$ (see Figure 14).

Each visual map $\mathcal{V}$ is composed of primitives $\mathcal{P}$, described by a parameter vector P. We dispose of an estimate $\widehat{P_0}$ of P and of an error covariance matrix W_{P_0}.

The coordinate frames $\mathcal{F}$ and $\mathcal{F}'$ are related by a rigid displacement $\mathcal{D}$ such that each point M' of $\mathcal{F}'$ is related to a point M of $\mathcal{F}$ by the relation

$$O'M' = ROM + t$$

where R is the rotation matrix and t the translation vector of the displacement $\mathcal{D}$. We also have an estimate $\widehat{D_0}$ of D, with an error covariance matrix W_{D_0}.

5.2 Defining geometric relations

We define a set of geometric relations between the primitives $\mathcal{P}$ and $\mathcal{P}'$ of two visual maps $\mathcal{V}$ and $\mathcal{V}'$. These relations are given in table 1.

This list of relations/primitives is not exhaustive but only demonstrative. The relation "identical" expresses the fact that the primitives $\mathcal{P}$ and $\mathcal{P}'$ represented in $\mathcal{V}$ and $\mathcal{V}'$ actually describe the same physical primitive. The relation "included" expresses that $\mathcal{P}$ describes a physical primitive which is part of the physical primitive described by $\mathcal{P}'$. The relations "parallel" and "orthogonal" are interpreted in a similar fashion.

Each geometric relation can be expressed by a vector equation of the form :

$$f_i(P, P', D) = 0 \tag{21}$$

relations	points	lines	planes
points	$\equiv$	$\subset$	$\subset$
lines		$\equiv // \perp$	$\subset // \perp$
planes			$\equiv // \perp$

Table 1: Relations between the primitives; $\equiv$: identical to; $\subset$: included in; $\|$: parallel to; $\perp$: orthogonal to.

5.3 Expliciting geometric relations

We now explicit equation 21 for the geometric relations of table 1. We denote by $\overline{P}$ the parameters of the primitive $\overline{\mathcal{P}} = D(\mathcal{P})$, the image of $\mathcal{P}$ by the rigid displacement D. The computation of $\overline{P}$ from P is, in the case of points, $\overline{OM} = R\,OM + t$. The case of lines and planes was explicited in the previous Section. The measurement equations are the following ones :

Point-Point:

$$\text{relation } \equiv \; : O'M' - \overline{OM} = 0$$

Point-Line: assuming the line is not orthogonal to the z-axis :

$$\text{relation } \subset \; : \begin{cases} \bar{x} - a'\bar{z} - p' &= 0 \\ \bar{y} - b'\bar{z} - q' &= 0 \end{cases}$$

Point-Plane: assuming the plane is not parallel to the z-axis :

$$\text{relation } \subset \; : a'\bar{x} + b'\bar{y} + \bar{z} + c' = 0$$

Line-Line: assuming the two lines are not orthogonal to the z-axis :

$$\begin{aligned} \text{relation } \equiv &: (a', b', c', d')^t - (\bar{a}, \bar{b}, \bar{c}, \bar{d})^t = 0 \\ \text{relation } \| &: (a', b')^t - (\bar{a}, \bar{b})^t = 0 \\ \text{relation } \perp &: a'\bar{a} + b'\bar{b} + 1 = 0 \end{aligned}$$

Line-Plane: assuming the line is not orthogonal and the plane not parallel to the z-axis :

$$\begin{aligned} \text{relation } \subset &: \begin{cases} a'\bar{a} + b'\bar{b} + 1 &= 0 \\ a'\bar{p} + b'\bar{q} + c' &= 0 \end{cases} \\ \text{relation } \| &: a'\bar{a} + b'\bar{b} + 1 = 0 \\ \text{relation } \perp &: (a', b')^t - (\bar{a}, \bar{b})^t = 0 \end{aligned}$$

Plane-Plane: assuming the plane not parallel to the z-axis :

$$\begin{aligned} \text{relation } \equiv &: (a', b', c')^t - (\bar{a}, \bar{b}, \bar{c})^t = 0 \\ \text{relation } \| &: (a', b')^t - (\bar{a}, \bar{b})^t = 0 \\ \text{relation } \perp &: a'\bar{a} + b'\bar{b} + 1 = 0 \end{aligned}$$

This approach should be compared to that of [Mun86].

5.4 Registration

The registration (or matching) of two primitives $\mathcal{P}$ and $\mathcal{P}'$ consists in detecting that their parameters P and P' verify equation 21 for one of the above listed geometric relations,with respect to the current noisy estimates $(\hat{P}_0, W_{P_0})$, $(\hat{P'}_0, W_{P'_0})$, and $(\hat{D}_0, W_{D_0})$ of P, P' and D.

This "detection" is done by computing between each pair of primitive the generalized Mahalanohis distance explicited by equation 11, and by matching a pair of primitive each time the χ^2 acceptation test given by inequality 12 is verified, i.e. when

$$d(\hat{P}_0, W_{P_0}, \hat{P'}_0, W_{P'_0}, \hat{D}_0, W_{D_0}) < \epsilon$$

In order to avoid a $0(n^2)$ complexity algorithm, it is of course possible to use additional control structures to select a subset of candidate primitives for each test. For instance, to test the relation "$\equiv$" between points or lines, bucketing techniques can be used with efficiency (see for instance [AFFT85]).

5.5 Motion

Having registered two primitives $\mathcal{P}$ and $\mathcal{P}'$, the motion problem consists in reducing the uncertainty on the *motion parameters* D while taking into account the uncertainty on the parameters P, P' and D.

This is done by setting $a = D$ and $x = (P, P')^t$, and by using the relation equation 21 as a measurement equation 1:

$$f_i(x, a) \equiv f_i((P, P'), D) = 0$$

Starting from the initial estimate $\hat{a}_0 = \hat{D}_0$, $S_0 = W_{D_0}$, and using the measurement $\hat{x}_1 = (\hat{P}_0, \hat{P'}_0)^t)$ with

$$W_1 = \begin{pmatrix} W_{P_0} & 0 \\ 0 & W_{P'_0} \end{pmatrix}$$

one applies the EKF formalism to obtain a new estimate $\hat{a}_1$ of the motion with a reduced covariance matrix $S_1 < S_0$. (In the sense $S_0 - S_1$ non-negative.)

This process is recursively repeated: at iteration i, if a new pair of primitives can be registered with the new motion estimate $(\hat{a}_{i-1}, S_{i-1})$, the additional measurement equations they bring lead

to a new better estimate $\hat{a}_i$ of the motion with a still reduced covariance matrix S_i. This process ends after the matching of k primitives with a final estimate $(\hat{a}_k,\ S_k)$ of the motion parameter D.

5.6 Fusion

5.6.1 General Fusion

The fusion problem is exactly the dual of the motion problem, as it consists, after the registration of 2 primitives, in reducing the uncertainty on the *primitive parameters* P and P' while taking into account the uncertainty on the parameters P, P' and D.

This is done by "switching the attention", i.e. by choosing $a = (P, P')^t$ and $x = D$ while using again the relation equation 21 as a measurement equation 1:

$$f_i(x, a) \equiv f_i(D, (P, P')) = 0$$

The initial estimate is taken as $\hat{a}_0 = (\hat{P}_0, \hat{P'}_0)^t$ and

$$S_0 = \begin{pmatrix} W_{P_0} & 0 \\ 0 & W_{P'_0} \end{pmatrix}$$

and one uses the measurement $\hat{x}_1 = \hat{D}_0$ with $W_1 = W_{D_0}$ to apply the EKF formalism and obtain a new estimate $\hat{a}_1$ of the primitive parameters with a reduced covariance matrix $S_1 < S_0$.

If additional relations hold between these primitives and other ones, the same treatment allows for a further reduction in their parameters uncertainty, and therefore a more accurate estimation of the primitive parameters.

5.6.2 Forgetting primitives

After the treatment of a constraint, the parameters P_1 and P'_1 of the primitives are usually correlated, which means that the covariance matrix

$$\mathrm{cov}(\hat{P}_1, \hat{P'}_1) = \begin{pmatrix} W_{P_1} & W_{P_1 P'_1} \\ W_{P'_1 P_1} & W_{P_1} \end{pmatrix}$$

contains $W_{P'_1 P_1} = W^t_{P_1 P'_1} \neq 0$.

Therefore, it is no longer possible to treat independently $\mathcal{P}$ and $\mathcal{P}'$ in successive measurement equations. One as to consider them as a new primitive, either by keeping only one of them, or the union of them.

For instance, if one updates the parameters of $\mathcal{P}'$ with those of an "identical" primitive $\mathcal{P}$ observed in a previous visual map, one keeps only the updated parameters of $\mathcal{P}'$ in the new map, with their covariance matrix $W_{p'}$, forgetting the previous paramaters $\mathcal{P}$ after having used them.

On the other hand, if one updates the parameters of two lines by detecting that they are orthogonal, one keeps the new primitive formed by the union of the updated two lines, with the corresponding covariance matrix. One must use this kind of relation carefully, in order to control the size of the state parameter a.

5.6.3 Autofusion

In the special case where $\mathcal{V} \equiv \mathcal{V}'$, all primitives come from the same visual map, and the motion parameters vanishes as they correspond to the identity transform and are perfectly known.

Nevertheless, one can still detect the previous geometric relations between pairs of primitives p and p', and use them to reduce the uncertainty on the primitives parameters.

6 Experimental Results

The basic principles presented in this paper were tested on a variety of synthetic and real data. The interested reader can find registration and motion results with real points and lines in [AF87b], registration and fusion results with synthetic and real points and lines in [AF87a] and results on the building of global 3D maps from passive stereovision in [AL87]. In this paper we only present results of the motion estimation from two 3D maps, of the fusion of several inaccurate 3D maps, and of the detection of colinearity within a single 3D map (what we called "autofusion"). In each of these examples, the 3D map is made of 3D lines.

6.1 Registration and Motion

Figure 15 shows the edges of a triplet of images taken by the mobile robot in a first position. From these edges, the trinocular stereovision system computes a set of 3D segments. Each 3D segment is represented by the parameters (a, b, p, q) of the 3D line supporting it and by the error covariance computed — as it is explained in section 4.2 — from the uncertainty on the edge points in the three images (we took an isotropic gaussian density function of covariance 1 pixel around each edge point). Each 3D line is bounded by 2 endpoints obtained from the endpoints measured in the 3 images which are projected on the reconstructed 3D line.

We show in figures 16 and 17 respectively the horizontal and vertical projections of the reconstructed 3D segments. We also show the uncertainty attached to the reconstructed 3D lines by showing the uncertainty it produces on the coordinates of their endpoints. The 95% confidence regions of the endpoints positions are ellipsoids whose projections are the ellipses shown in figures 16 and 17. One can see the anisotropic distribution of the uncertainty on the three coordinates of the points and its variation as a function of their position relative to the cameras (which are located grossly in the middle of the front view and at the bottom of the top view).

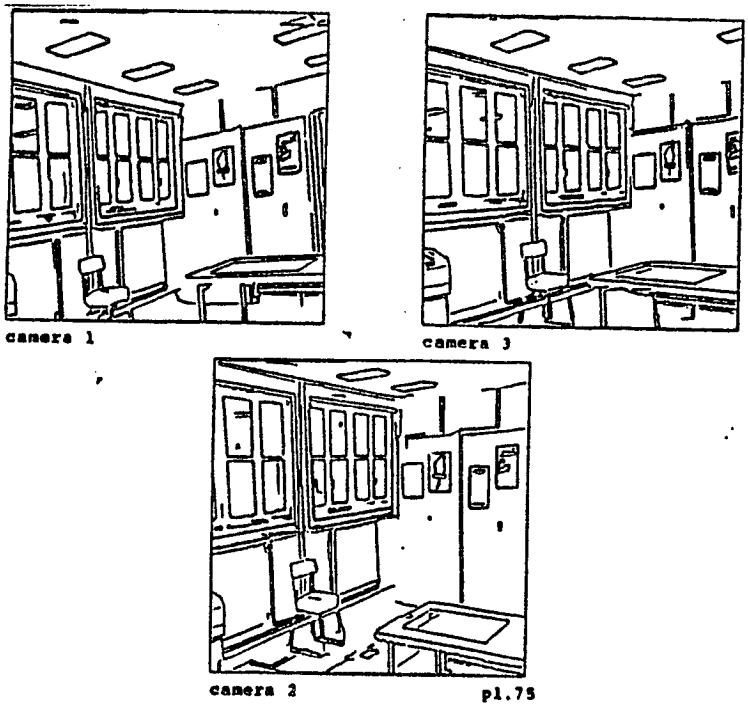

Figure 15: Triplet of images taken in position 1

The robot now moves a little, a new triplet of images is taken (figure 18) and another set of 3D lines is computed. Initially, the robot is given a very crude estimate of its motion between the two views. Applying this crude estimate to the 3D lines obtained in position 1, and projecting them in one of the images obtained in position 2 (the image of camera 3), one obtains the crude superimposition observed in figure 19. Solid lines are the transformed 3D segments computed in position 1, while the dotted lines are the 2D segments observed in position 2.

We now ask the system to discover the relation "$\equiv$" between the 3D lines (see section 5.3) reconstructed in positions 1 and 2, given the initial crude motion estimate and its uncertainty. The program takes each 3D line in position 1, applies the noisy current motion estimate to place it in the 3D map obtained in position 2 with a new covariance matrix (combining the initial uncertainty with the motion uncertainty), and computes its Mahalanobis distance (equation 11) to all the other lines of position 2.

The program detects a match each time a pair of lines passes the χ^2 test of equation 12. If a line in position 1 can be matched to several lines in position 2, this is an *ambiguous* match, and nothing is done. On the other hand, each time an *unambiguous* match is found, the parameters of the motion are updated as it is explained in section 5.5. As the uncertainty on motion decreases after each new match, some previously ambiguous matches can now become unambiguous. Therefore the entire matching process is repeated until no more line can be matched (3 iterations in this example). The final estimate of the motion is very accurate, as can be seen in figure 20 where the obtained superimposition is now almost perfect.

Applying exactly the same technique to a set of 6 triplets of views taken during the motion of

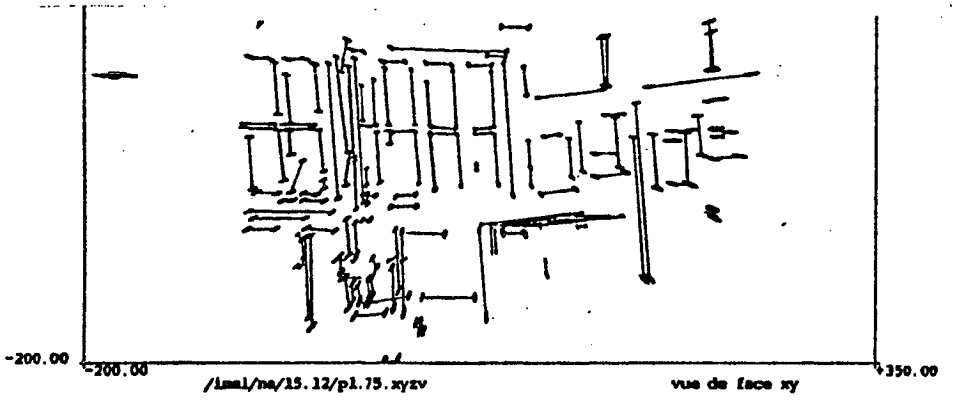

Figure 16: Front view of reconstructed 3D lines

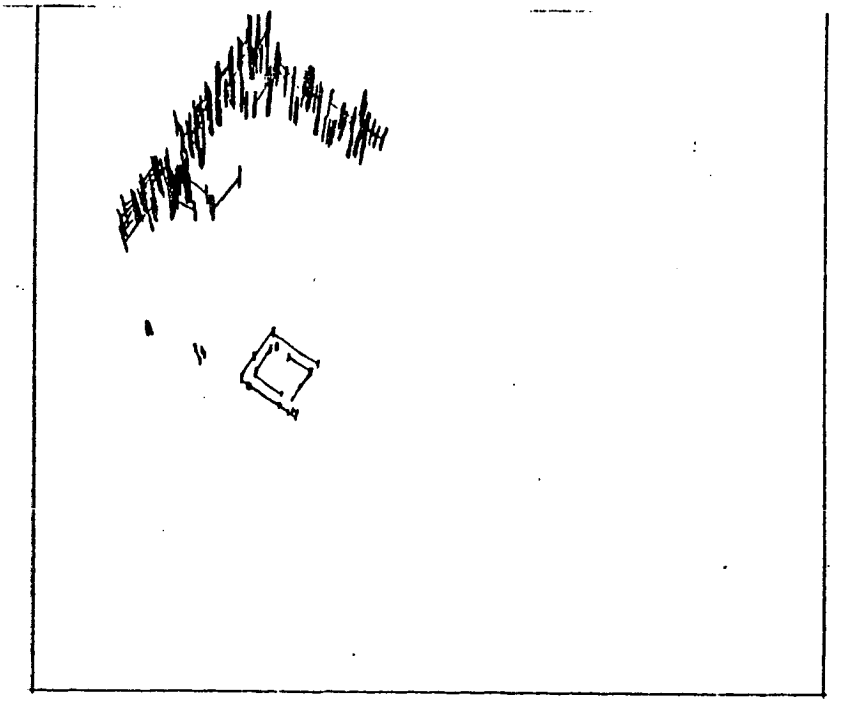

Figure 17: Top view of reconstructed 3D lines

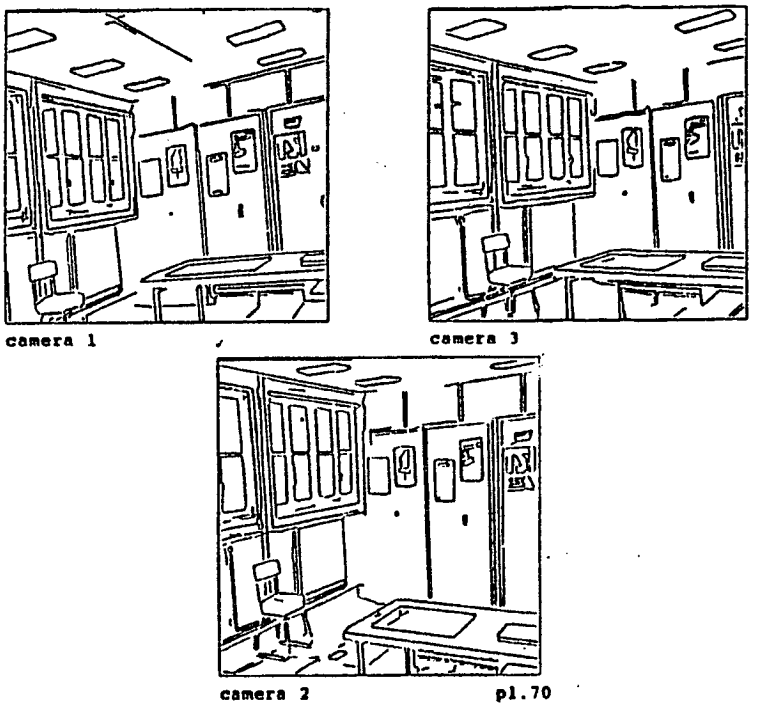

Figure 18: Triplet of images taken in position 2

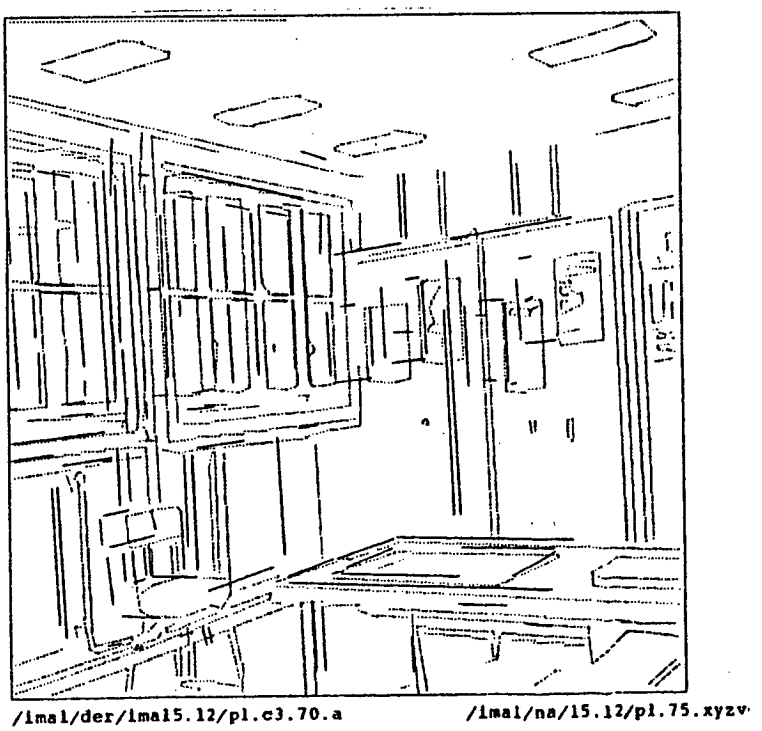

Figure 19: Superimposition of 3D segments of position 1 with 2D edges of position 2 (crude initial motion estimate)

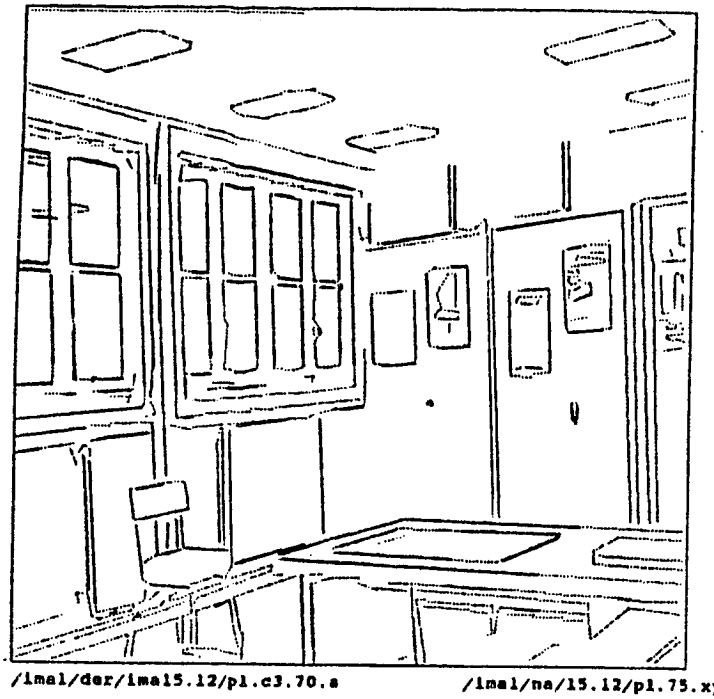

Figure 20: Superimposition of 3D segments of position 1 with 2D edges of position 2 (final motion estimate)

the robot, the sytem was able to build a global 3D map of the room (figure 21) where rotating segments at the bottom right are the computed successive robot positions.

6.2 Registration, Motion and Fusion

In this experiment, the robot is looking from 4 different positions at a regular pattern (figures 22 and 23) formed by vertical lines floating in front of horizontal lines, and builds in each position a local 3D map. Exactly the same technique as in the previous example was used to register each successive local 3D map, and put all of them in a single absolute reference frame. Figure 24 shows the resulting 3D map before fusion. Fusion is achieved by discovering the relation "$\equiv$" computed between lines in the global 3D map, and taking into account the uncertainty on the 3D lines due to their reconstruction and to the successive motion estimations. Fusion yields a reduction from 1808 to 650 segments and improved accuracy, as can be seen by looking at the front and top view of the reconstructed 3D pattern after fusion (Figure 25).

6.3 Detecting colinearity in space

In this experiment, the robot is looking at the regular pattern only once. We show in figure 26 the vertical and horizontal projections of the initially reconstructed 3D segments. We also show the uncertainty attached to reconstructed 3D lines by showing the uncertainty it produces on the coordinates of their endpoints. (in the same way as in the first experiment).

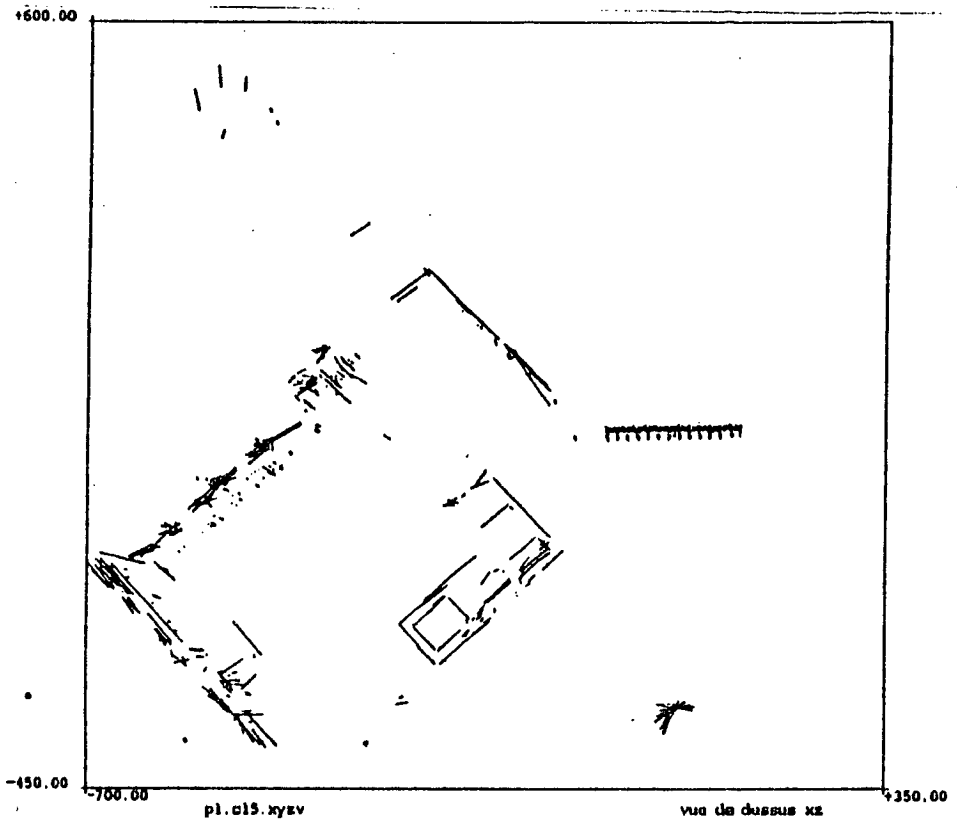

Figure 21: Top view of a global 3D map of the room computed from 6 local 3D maps

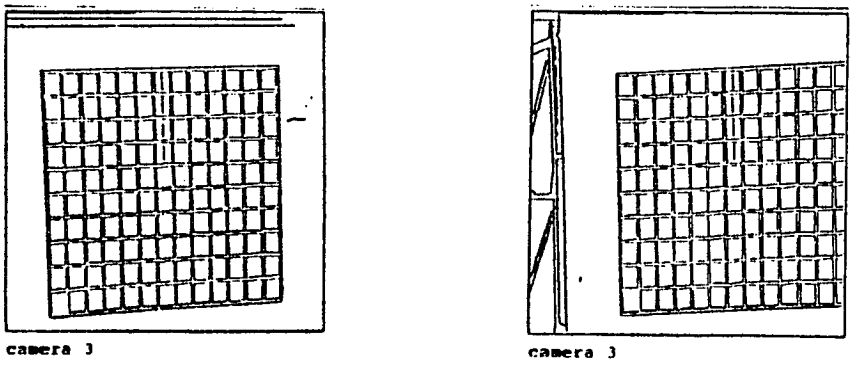

Figure 22: A regular grid observed from positions 1 and 2

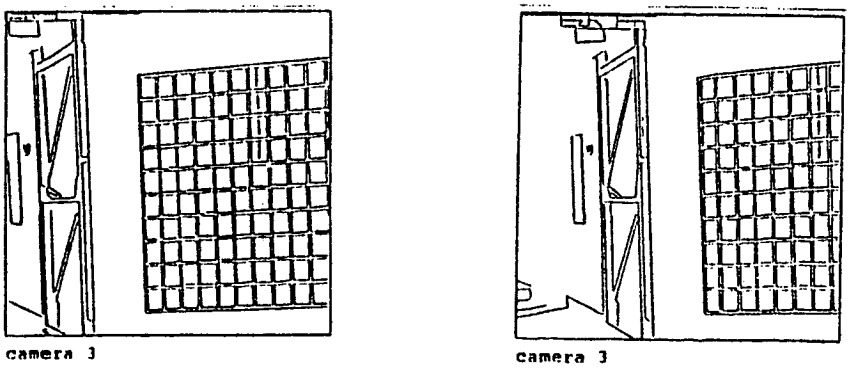

Figure 23: A regular grid observed from positions 3 and 4

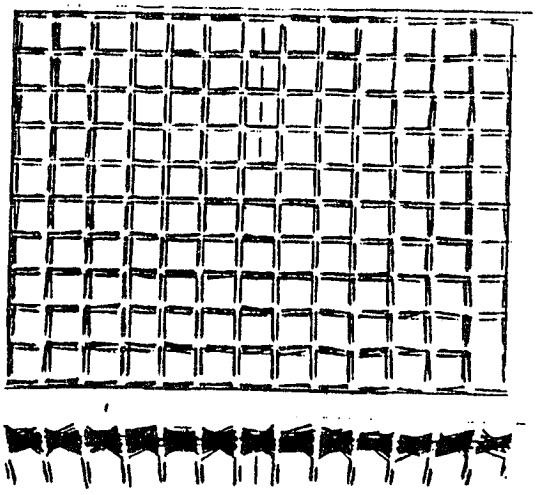

Figure 24: Front and top view of reconstructed 3D lines before identical lines are detected

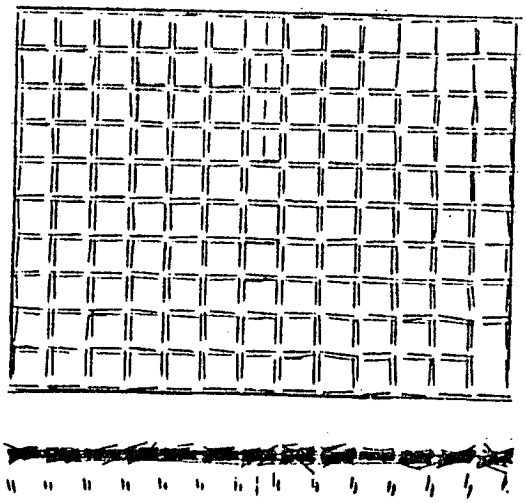

Figure 25: Front and top view of 3D lines after the fusion of identical lines

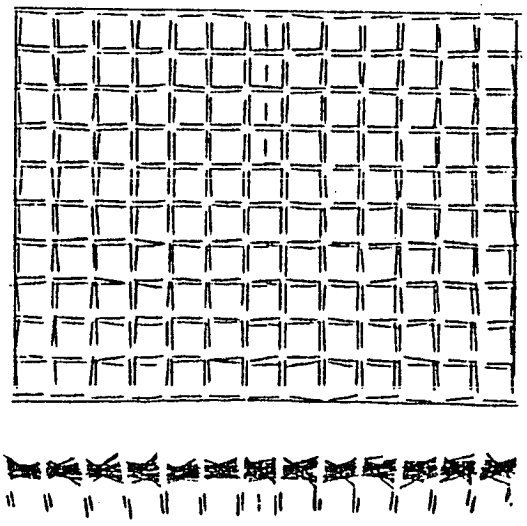

Figure 26: Front and top view of reconstructed 3D lines, before colinearity is detected

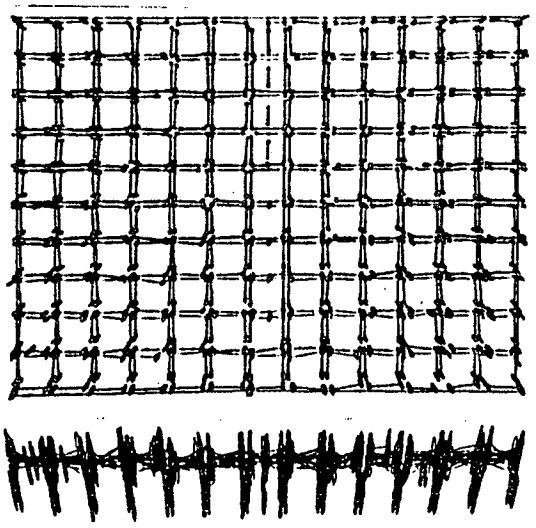

Figure 27: Initial uncertainty attached to 3D lines endpoints

We now ask the system to discover the relation "$\equiv$" between the 3D lines (see section 5.3). The program takes a first 3D line, computes its Mahalanobis distance (equation 11) to all the other lines of the scene, and accepts the first line which passes the χ^2 test of equation 12. The two lines are fused using the technique of section 5.6 and one keeps only the parameters of the optimal line representing both of them with an updated covariance matrix. The remaining lines are now compared to this new virtual line still with the Mahalanobis distance of equation 11 but with the new updated covariance matrix, while the χ^2 test of equation 12 remains unchanged. This process is repeated until no more line can be matched with the first one, and then repeated with all the remaining unmatched lines.

The result is a reduced set of virtual lines on which the endpoints of the original segments have been projected, as shown in figure 28. The uncertainty on the line parameters has been greatly reduced: figure 29 shows the resulting uncertainty on the lines endpoints, which agrees very well with the reality.

References

[AF87a] N. Ayache and O.D. Faugeras. Building a consistent 3d representation of a mobile robot environment by combining multiple stereo views. In *Proc. International Joint Conference on Artificial Intelligence*, August 1987. Milano, Italy.

[AF87b] N. Ayache and O.D. Faugeras. Building, registrating and fusing noisy visual maps. In *Proc. First International Conference on Computer Vision*, pages 73–82, IEEE, June 1987. London, U.K., also an INRIA Internal Report 596, 1986.

[AF87c] N. Ayache and B. Faverjon. Efficient registration of stereo images by matching graph descriptions of edge segments. *The International Journal of Computer Vision*, 1(2),

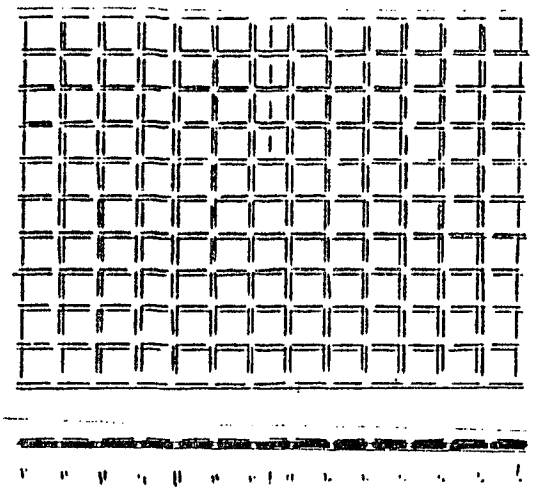

Figure 28: Front and top view of 3D lines when colinearity is discovered and enforced

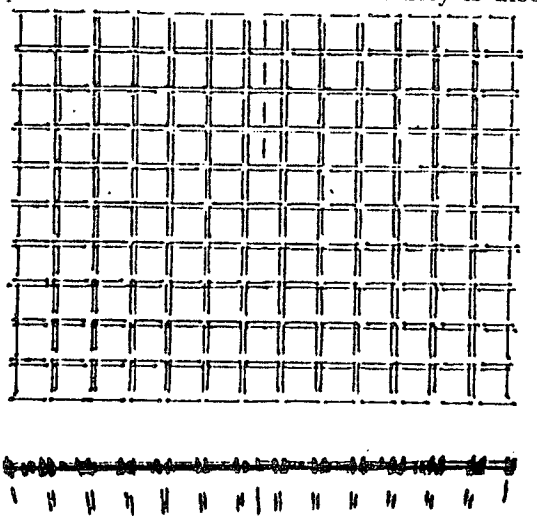

Figure 29: Uncertainty attached to 3D lines endpoints after the fusion of colinear segments

April 1987.

[AFFT85] N. Ayache, O.D. Faugeras, B. Faverjon, and G. Toscani. Matching depth maps obtained by passive stereovision. In *Proceedings of Third Workshop on Computer Vision: Representation and Control*, pages 197–204, IEEE, October 1985.

[AL87] N. Ayache and F. Lustman. Fast and reliable passive trinocular stereovision. In *Proc. First International Conference on Computer Vision*, pages 422–427, IEEE, June 1987. London, U.K.

[BC86] Bolle and Cooper. On optimally combining pieces of information, with application to estimating 3d complex-object position from range data. In *IEEE Transactions on PAMI*, pages 619–638, 1986.

[Can86] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8 No6:679–698, 1986.

[Der87] R. Deriche. Using canny's criteria to derive an optimal edge detector recursively implemented. *The International Journal of Computer Vision*, 2, April 1987.

[Dur86] H.F. Durrant-Whyte. Consistent integration and propagation of disparate sensor observations. In *Proc. International Conference on Robotics and Automation*, pages 1464–1469, April 1986. San Francisco, CA, USA.

[FAF86] O.D. Faugeras, N. Ayache, and B. Faverjon. Building visual maps by combining noisy stereo measurements. In *Proc. International Conference on Robotics and Automation*, pages 1433–1438, April 1986. San Francisco, CA, USA.

[FH86] O.D. Faugeras and M. Hebert. The representation, recognition, and locating of 3d shapes from range data. *The International Journal of Robotics Research*, 5, No 3:27–52, 1986.

[FLT87] O.D. Faugeras, F. Lustman, and G. Toscani. Motion and structure from motion from point and line matches. In *Proc. First International Conference on Computer Vision*, pages 25–34, IEEE, June 1987. London, U.K.

[HM87] Hager and M. Mintz. *Estimation Procedures for Robust Sensor Control, in the integration of sensing with actuation to form a robust intelligent control system.* GRASPLAB Report 97, Department of Computer and Information Science, Moore School, University of Pennsylvania, March 1987.

[Jaz70] A.M. Jazwinsky. *Stochastic Processes and Filtering Theory.* Academic Press, 1970.

[Kan87] T. Kanade. *Three-Dimensional Machine Vision.* Kluwer Academic Publishers, 1987.

[KTT87] D.J. Kriegman, E. Triendl, and Binford T.O. A mobile robot: sensing, planning and locomotion. In *Proc. International Conference on Robotics and Automation*, pages 402–408, IEEE, 1987. Raleigh,North Carolina.

[MM87] McKendall and M. Mintz. *Models of Sensor Noise and Optimal Algorithms for Estimation and Quantization in Vision Systems.* GRASPLAB Report 97, Department of Computer and Information Science, Moore School, University of Pennsylvania, March 1987.

[MS86] Matthies and Shafer. *Error modelling in Stereo Navigation.* Technical Report, University of Carnegie-Mellon, Department of Computer Science, 1986. pp. 86–140.

[Mun86] J.L. Mundy. Reasoning about 3-d space with algebraic deduction. In O. D. Faugeras and Georges Giralt, editors, *Robotics Research, The Third International Symposium*, pages 117–124, MIT Press, 1986.

[P*87] J. Porrill et al. Optimal combination and constraints for geometrical sensor data. March 1987. to appear.

[SC87] R.C. Smith and P. Cheeseman. On the representation and estimation of spatial uncertainty. *International Journal of Robotics Research*, 5(4):56–68, 1987.

Imprimé en France
par
l'Institut National de Recherche en Informatique et en Automatique